

ECGR_HW4

March 29, 2023

1 ECGR HW 4

Patrick Flynn | 801055057

Set everything up:

```
[1]: !pip install torch torchvision
      !pip install d2l==1.0.0a1.post0
      !pip install matplotlib_inline
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Requirement already satisfied: torch in /usr/local/lib/python3.9/dist-packages
(1.13.1+cu116)
Requirement already satisfied: torchvision in /usr/local/lib/python3.9/dist-
packages (0.14.1+cu116)
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.9/dist-packages (from torch) (4.5.0)
Requirement already satisfied: requests in /usr/local/lib/python3.9/dist-
packages (from torchvision) (2.27.1)
Requirement already satisfied: numpy in /usr/local/lib/python3.9/dist-packages
(from torchvision) (1.22.4)
Requirement already satisfied: pillow!=8.3.*,>=5.3.0 in
/usr/local/lib/python3.9/dist-packages (from torchvision) (8.4.0)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in
/usr/local/lib/python3.9/dist-packages (from requests->torchvision) (1.26.15)
Requirement already satisfied: charset-normalizer~=2.0.0 in
/usr/local/lib/python3.9/dist-packages (from requests->torchvision) (2.0.12)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.9/dist-packages (from requests->torchvision) (2022.12.7)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.9/dist-
packages (from requests->torchvision) (3.4)
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting d2l==1.0.0a1.post0
  Downloading d2l-1.0.0a1.post0-py3-none-any.whl (93 kB)
                                93.0/93.0 KB
1.7 MB/s eta 0:00:00
Requirement already satisfied: pandas in /usr/local/lib/python3.9/dist-
```

```

packages (from d2l==1.0.0a1.post0) (1.4.4)
Requirement already satisfied: gym in /usr/local/lib/python3.9/dist-packages
(from d2l==1.0.0a1.post0) (0.25.2)
Collecting matplotlib-inline
  Downloading matplotlib_inline-0.1.6-py3-none-any.whl (9.4 kB)
Collecting jupyter
  Downloading jupyter-1.0.0-py2.py3-none-any.whl (2.7 kB)
Requirement already satisfied: numpy in /usr/local/lib/python3.9/dist-packages
(from d2l==1.0.0a1.post0) (1.22.4)
Requirement already satisfied: requests in /usr/local/lib/python3.9/dist-
packages (from d2l==1.0.0a1.post0) (2.27.1)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.9/dist-
packages (from d2l==1.0.0a1.post0) (3.7.1)
Requirement already satisfied: importlib-metadata>=4.8.0 in
/usr/local/lib/python3.9/dist-packages (from gym->d2l==1.0.0a1.post0) (6.1.0)
Requirement already satisfied: gym-notices>=0.0.4 in
/usr/local/lib/python3.9/dist-packages (from gym->d2l==1.0.0a1.post0) (0.0.8)
Requirement already satisfied: cloudpickle>=1.2.0 in
/usr/local/lib/python3.9/dist-packages (from gym->d2l==1.0.0a1.post0) (2.2.1)
Requirement already satisfied: notebook in /usr/local/lib/python3.9/dist-
packages (from jupyter->d2l==1.0.0a1.post0) (6.3.0)
Requirement already satisfied: nbconvert in /usr/local/lib/python3.9/dist-
packages (from jupyter->d2l==1.0.0a1.post0) (6.5.4)
Collecting qtconsole
  Downloading qtconsole-5.4.1-py3-none-any.whl (120 kB)
120.9/120.9

```

KB 7.4 MB/s eta 0:00:00

```

Requirement already satisfied: ipykernel in /usr/local/lib/python3.9/dist-
packages (from jupyter->d2l==1.0.0a1.post0) (5.3.4)
Requirement already satisfied: jupyter-console in /usr/local/lib/python3.9/dist-
packages (from jupyter->d2l==1.0.0a1.post0) (6.1.0)
Requirement already satisfied: ipywidgets in /usr/local/lib/python3.9/dist-
packages (from jupyter->d2l==1.0.0a1.post0) (7.7.1)
Requirement already satisfied: fonttools>=4.22.0 in
/usr/local/lib/python3.9/dist-packages (from matplotlib->d2l==1.0.0a1.post0)
(4.39.2)
Requirement already satisfied: importlib-resources>=3.2.0 in
/usr/local/lib/python3.9/dist-packages (from matplotlib->d2l==1.0.0a1.post0)
(5.12.0)
Requirement already satisfied: pillow>=6.2.0 in /usr/local/lib/python3.9/dist-
packages (from matplotlib->d2l==1.0.0a1.post0) (8.4.0)
Requirement already satisfied: python-dateutil>=2.7 in
/usr/local/lib/python3.9/dist-packages (from matplotlib->d2l==1.0.0a1.post0)
(2.8.2)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.9/dist-
packages (from matplotlib->d2l==1.0.0a1.post0) (0.11.0)
Requirement already satisfied: kiwisolver>=1.0.1 in

```

/usr/local/lib/python3.9/dist-packages (from matplotlib->d2l==1.0.0a1.post0)
 (1.4.4)
 Requirement already satisfied: pyparsing>=2.3.1 in
 /usr/local/lib/python3.9/dist-packages (from matplotlib->d2l==1.0.0a1.post0)
 (3.0.9)
 Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.9/dist-
 packages (from matplotlib->d2l==1.0.0a1.post0) (23.0)
 Requirement already satisfied: contourpy>=1.0.1 in
 /usr/local/lib/python3.9/dist-packages (from matplotlib->d2l==1.0.0a1.post0)
 (1.0.7)
 Requirement already satisfied: traitlets in /usr/local/lib/python3.9/dist-
 packages (from matplotlib-inline->d2l==1.0.0a1.post0) (5.7.1)
 Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.9/dist-
 packages (from pandas->d2l==1.0.0a1.post0) (2022.7.1)
 Requirement already satisfied: urllib3<1.27,>=1.21.1 in
 /usr/local/lib/python3.9/dist-packages (from requests->d2l==1.0.0a1.post0)
 (1.26.15)
 Requirement already satisfied: charset-normalizer~=2.0.0 in
 /usr/local/lib/python3.9/dist-packages (from requests->d2l==1.0.0a1.post0)
 (2.0.12)
 Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.9/dist-
 packages (from requests->d2l==1.0.0a1.post0) (3.4)
 Requirement already satisfied: certifi>=2017.4.17 in
 /usr/local/lib/python3.9/dist-packages (from requests->d2l==1.0.0a1.post0)
 (2022.12.7)
 Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.9/dist-
 packages (from importlib-metadata>=4.8.0->gym->d2l==1.0.0a1.post0) (3.15.0)
 Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.9/dist-
 packages (from python-dateutil>=2.7->matplotlib->d2l==1.0.0a1.post0) (1.16.0)
 Requirement already satisfied: ipython>=5.0.0 in /usr/local/lib/python3.9/dist-
 packages (from ipykernel->jupyter->d2l==1.0.0a1.post0) (7.9.0)
 Requirement already satisfied: jupyter-client in /usr/local/lib/python3.9/dist-
 packages (from ipykernel->jupyter->d2l==1.0.0a1.post0) (6.1.12)
 Requirement already satisfied: tornado>=4.2 in /usr/local/lib/python3.9/dist-
 packages (from ipykernel->jupyter->d2l==1.0.0a1.post0) (6.2)
 Requirement already satisfied: widgetsnbextension~=3.6.0 in
 /usr/local/lib/python3.9/dist-packages (from
 ipywidgets->jupyter->d2l==1.0.0a1.post0) (3.6.3)
 Requirement already satisfied: ipython-genutils~=0.2.0 in
 /usr/local/lib/python3.9/dist-packages (from
 ipywidgets->jupyter->d2l==1.0.0a1.post0) (0.2.0)
 Requirement already satisfied: jupyterlab-widgets>=1.0.0 in
 /usr/local/lib/python3.9/dist-packages (from
 ipywidgets->jupyter->d2l==1.0.0a1.post0) (3.0.6)
 Requirement already satisfied: pygments in /usr/local/lib/python3.9/dist-
 packages (from jupyter-console->jupyter->d2l==1.0.0a1.post0) (2.14.0)
 Requirement already satisfied: prompt-toolkit!=3.0.0,!3.0.1,<3.1.0,>=2.0.0 in
 /usr/local/lib/python3.9/dist-packages (from jupyter-

```

console->jupyter->d2l==1.0.0a1.post0) (2.0.10)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.9/dist-
packages (from nbconvert->jupyter->d2l==1.0.0a1.post0) (2.1.2)
Requirement already satisfied: tinycss2 in /usr/local/lib/python3.9/dist-
packages (from nbconvert->jupyter->d2l==1.0.0a1.post0) (1.2.1)
Requirement already satisfied: pandocfilters>=1.4.1 in
/usr/local/lib/python3.9/dist-packages (from
nbconvert->jupyter->d2l==1.0.0a1.post0) (1.5.0)
Requirement already satisfied: Jinja2>=3.0 in /usr/local/lib/python3.9/dist-
packages (from nbconvert->jupyter->d2l==1.0.0a1.post0) (3.1.2)
Requirement already satisfied: nbclient>=0.5.0 in /usr/local/lib/python3.9/dist-
packages (from nbconvert->jupyter->d2l==1.0.0a1.post0) (0.7.2)
Requirement already satisfied: entrypoints>=0.2.2 in
/usr/local/lib/python3.9/dist-packages (from
nbconvert->jupyter->d2l==1.0.0a1.post0) (0.4)
Requirement already satisfied: bleach in /usr/local/lib/python3.9/dist-packages
(from nbconvert->jupyter->d2l==1.0.0a1.post0) (6.0.0)
Requirement already satisfied: mistune<2,>=0.8.1 in
/usr/local/lib/python3.9/dist-packages (from
nbconvert->jupyter->d2l==1.0.0a1.post0) (0.8.4)
Requirement already satisfied: beautifulsoup4 in /usr/local/lib/python3.9/dist-
packages (from nbconvert->jupyter->d2l==1.0.0a1.post0) (4.11.2)
Requirement already satisfied: defusedxml in /usr/local/lib/python3.9/dist-
packages (from nbconvert->jupyter->d2l==1.0.0a1.post0) (0.7.1)
Requirement already satisfied: jupyterlab-pygments in
/usr/local/lib/python3.9/dist-packages (from
nbconvert->jupyter->d2l==1.0.0a1.post0) (0.2.2)
Requirement already satisfied: jupyter-core>=4.7 in
/usr/local/lib/python3.9/dist-packages (from
nbconvert->jupyter->d2l==1.0.0a1.post0) (5.3.0)
Requirement already satisfied: lxml in /usr/local/lib/python3.9/dist-packages
(from nbconvert->jupyter->d2l==1.0.0a1.post0) (4.9.2)
Requirement already satisfied: nbformat>=5.1 in /usr/local/lib/python3.9/dist-
packages (from nbconvert->jupyter->d2l==1.0.0a1.post0) (5.8.0)
Requirement already satisfied: terminado>=0.8.3 in
/usr/local/lib/python3.9/dist-packages (from
notebook->jupyter->d2l==1.0.0a1.post0) (0.17.1)
Requirement already satisfied: Send2Trash>=1.5.0 in
/usr/local/lib/python3.9/dist-packages (from
notebook->jupyter->d2l==1.0.0a1.post0) (1.8.0)
Requirement already satisfied: pyzmq>=17 in /usr/local/lib/python3.9/dist-
packages (from notebook->jupyter->d2l==1.0.0a1.post0) (23.2.1)
Requirement already satisfied: argon2-cffi in /usr/local/lib/python3.9/dist-
packages (from notebook->jupyter->d2l==1.0.0a1.post0) (21.3.0)
Requirement already satisfied: prometheus-client in
/usr/local/lib/python3.9/dist-packages (from
notebook->jupyter->d2l==1.0.0a1.post0) (0.16.0)
Collecting qtpy>=2.0.1

```

```

    Downloading QtPy-2.3.1-py3-none-any.whl (84 kB)
      84.9/84.9 KB
5.6 MB/s eta 0:00:00
Requirement already satisfied: setuptools>=18.5 in
/usr/local/lib/python3.9/dist-packages (from
ipython>=5.0.0->ipykernel->jupyter->d2l==1.0.0a1.post0) (67.6.0)
Collecting jedi>=0.10
  Downloading jedi-0.18.2-py2.py3-none-any.whl (1.6 MB)
    1.6/1.6 MB
15.4 MB/s eta 0:00:00
Requirement already satisfied: pexpect in /usr/local/lib/python3.9/dist-
packages (from ipython>=5.0.0->ipykernel->jupyter->d2l==1.0.0a1.post0) (4.8.0)
Requirement already satisfied: backcall in /usr/local/lib/python3.9/dist-
packages (from ipython>=5.0.0->ipykernel->jupyter->d2l==1.0.0a1.post0) (0.2.0)
Requirement already satisfied: decorator in /usr/local/lib/python3.9/dist-
packages (from ipython>=5.0.0->ipykernel->jupyter->d2l==1.0.0a1.post0) (4.4.2)
Requirement already satisfied: pickleshare in /usr/local/lib/python3.9/dist-
packages (from ipython>=5.0.0->ipykernel->jupyter->d2l==1.0.0a1.post0) (0.7.5)
Requirement already satisfied: platformdirs>=2.5 in
/usr/local/lib/python3.9/dist-packages (from jupyter-
core>=4.7->nbconvert->jupyter->d2l==1.0.0a1.post0) (3.1.1)
Requirement already satisfied: fastjsonschema in /usr/local/lib/python3.9/dist-
packages (from nbformat>=5.1->nbconvert->jupyter->d2l==1.0.0a1.post0) (2.16.3)
Requirement already satisfied: jsonschema>=2.6 in /usr/local/lib/python3.9/dist-
packages (from nbformat>=5.1->nbconvert->jupyter->d2l==1.0.0a1.post0) (4.3.3)
Requirement already satisfied: wcwidth in /usr/local/lib/python3.9/dist-packages
(from prompt-toolkit!=3.0.0,!<3.0.1,<3.1.0,>=2.0.0->jupyter-
console->jupyter->d2l==1.0.0a1.post0) (0.2.6)
Requirement already satisfied: ptyprocess in /usr/local/lib/python3.9/dist-
packages (from terminado>=0.8.3->notebook->jupyter->d2l==1.0.0a1.post0) (0.7.0)
Requirement already satisfied: argon2-cffi-bindings in
/usr/local/lib/python3.9/dist-packages (from
argon2-cffi->notebook->jupyter->d2l==1.0.0a1.post0) (21.2.0)
Requirement already satisfied: soupsieve>1.2 in /usr/local/lib/python3.9/dist-
packages (from beautifulsoup4->nbconvert->jupyter->d2l==1.0.0a1.post0) (2.4)
Requirement already satisfied: webencodings in /usr/local/lib/python3.9/dist-
packages (from bleach->nbconvert->jupyter->d2l==1.0.0a1.post0) (0.5.1)
Requirement already satisfied: parso<0.9.0,>=0.8.0 in
/usr/local/lib/python3.9/dist-packages (from
jedi>=0.10->ipython>=5.0.0->ipykernel->jupyter->d2l==1.0.0a1.post0) (0.8.3)
Requirement already satisfied: pyparsing!=0.17.0,!<0.17.1,!<0.17.2,>=0.14.0 in
/usr/local/lib/python3.9/dist-packages (from
jsonschema>=2.6->nbformat>=5.1->nbconvert->jupyter->d2l==1.0.0a1.post0) (0.19.3)
Requirement already satisfied: attrs>=17.4.0 in /usr/local/lib/python3.9/dist-
packages (from
jsonschema>=2.6->nbformat>=5.1->nbconvert->jupyter->d2l==1.0.0a1.post0) (22.2.0)
Requirement already satisfied: cffi>=1.0.1 in /usr/local/lib/python3.9/dist-
packages (from argon2-cffi-

```

```
bindings->argon2-cffi->notebook->jupyter->d2l==1.0.0a1.post0) (1.15.1)
Requirement already satisfied: pycparser in /usr/local/lib/python3.9/dist-
packages (from cffi>=1.0.1->argon2-cffi-
bindings->argon2-cffi->notebook->jupyter->d2l==1.0.0a1.post0) (2.21)
Installing collected packages: qtpy, matplotlib-inline, jedi, qtconsole,
jupyter, d2l
Successfully installed d2l-1.0.0a1.post0 jedi-0.18.2 jupyter-1.0.0 matplotlib-
inline-0.1.6 qtconsole-5.4.1 qtpy-2.3.1
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Requirement already satisfied: matplotlib-inline in
/usr/local/lib/python3.9/dist-packages (0.1.6)
Requirement already satisfied: traitlets in /usr/local/lib/python3.9/dist-
packages (from matplotlib-inline) (5.7.1)
```

```
[2]: %matplotlib inline
import time
import torch
import torchvision
from torchvision import transforms
from d2l import torch as d2l
from torch import nn
import torch.nn.functional as F

d2l.use_svg_display()
```

1.1 Problem 1

1. Use the GRU example, adjust the hyperparameters (fully connected network and the number of hidden states) and analyze their influence on running time, perplexity, training and validation loss, and the output sequence (try a few examples).

The GRU

```
[22]: class GRU(d2l.RNN):
        def __init__(self, num_inputs, num_hiddens):
            d2l.Module.__init__(self)
            self.save_hyperparameters()
            self.rnn = nn.GRU(num_inputs, num_hiddens)
```

```
[7]: data = d2l.TimeMachine(batch_size=1024, num_steps=32)
```

Downloading ../data/timemachine.txt from
<http://d2l-data.s3-accelerate.amazonaws.com/timemachine.txt...>

Base line trained:

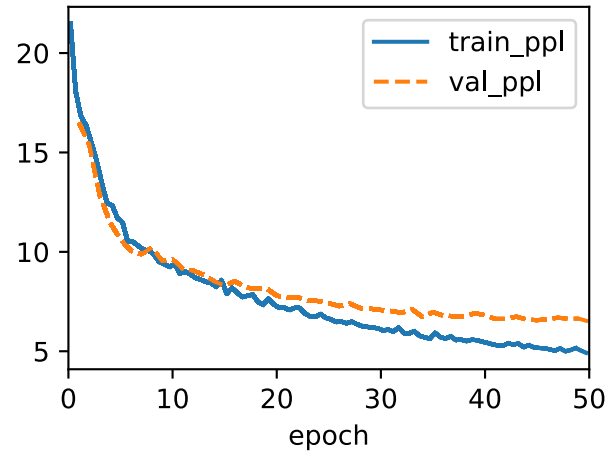
```
[23]: gru = GRU(num_inputs=len(data.vocab), num_hiddens=32)
model = d2l.RNNLM(gru, vocab_size=len(data.vocab), lr=4)
```

```

trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())

```

[23]: 'it has a conter and the ti'



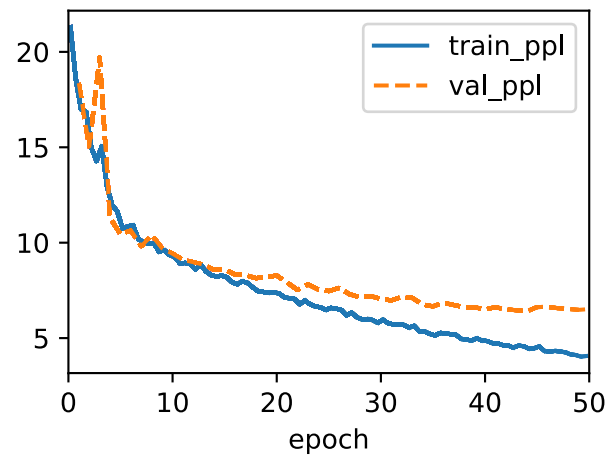
Adjusted 1:

```

[24]: gru = GRU(num_inputs=len(data.vocab), num_hiddens=64)
model = d2l.RNNLM(gru, vocab_size=len(data.vocab), lr=4)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())

```

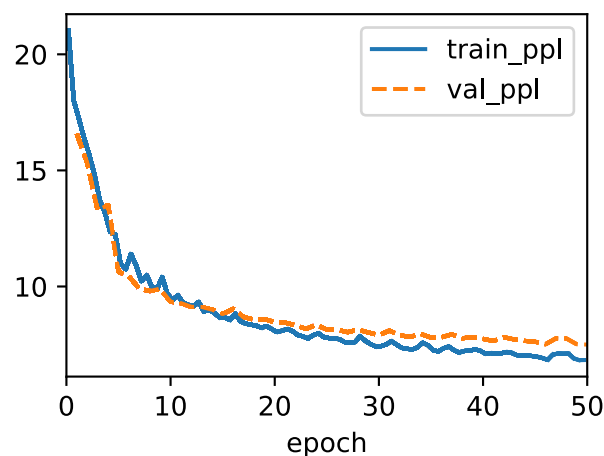
[24]: 'it has in the time travell'



Adjusted 2:

```
[25]: gru = GRU(num_inputs=len(data.vocab), num_hiddens=12)
model = d2l.RNNLM(gru, vocab_size=len(data.vocab), lr=4)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

```
[25]: 'it has the the the the the'
```



2. Use the LSTM example, adjust the hyperparameters (fully connected network and the number of hidden states) and analyze their influence on running time, perplexity, training and validation loss, and the output sequence (try a few examples)

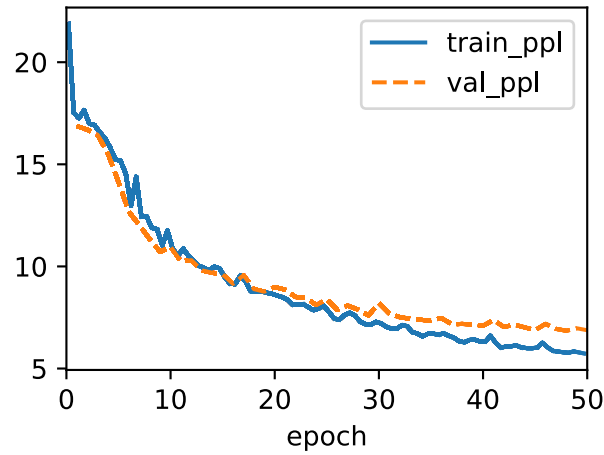
Base LSTM

```
[26]: class LSTM(d2l.RNN):
    def __init__(self, num_inputs, num_hiddens):
        d2l.Module.__init__(self)
        self.save_hyperparameters()
        self.rnn = nn.LSTM(num_inputs, num_hiddens)

    def forward(self, inputs, H_C=None):
        return self.rnn(inputs, H_C)
```

```
[27]: lstm = LSTM(num_inputs=len(data.vocab), num_hiddens=32)
model = d2l.RNNLM(lstm, vocab_size=len(data.vocab), lr=4)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

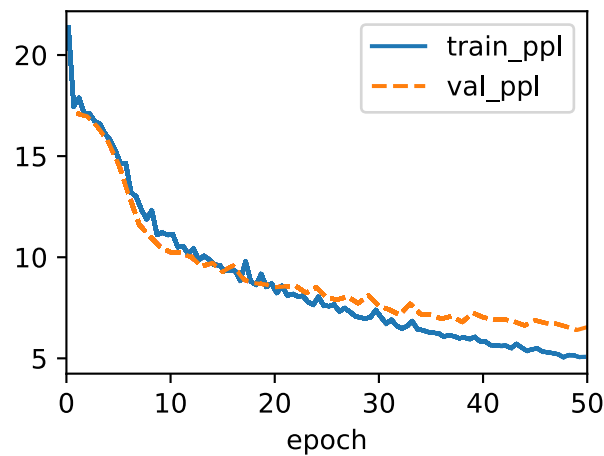

[27]: 'it has and the the that th'



Adjusted 1:

```
[28]: lstm = LSTM(num_inputs=len(data.vocab), num_hidden=64)
model = d2l.RNNLM(lstm, vocab_size=len(data.vocab), lr=4)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

[28]: 'it has the time traveller '



Adjusted 2:

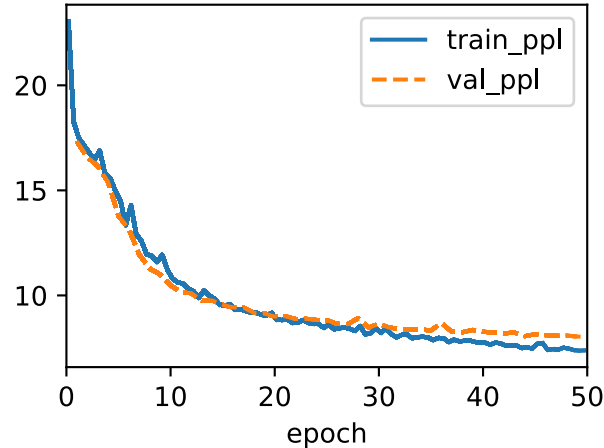
```
[29]: lstm = LSTM(num_inputs=len(data.vocab), num_hidden=12)
model = d2l.RNNLM(lstm, vocab_size=len(data.vocab), lr=4)
```

```

trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())

```

[29]: 'it has the the the the the'



3. Compare runtime for training and inference, computational and mode size complexities, training and validation loss, and the output sequence (try a few examples) for rnn.RNN, rnn.LSTM and rnn.GRU implementations with each other use the same hyperparameters for your comparison.

RNN implementation (for comparison)

```

[30]: class RNN(d2l.Module):
        def __init__(self, num_inputs, num_hiddens):
            super().__init__()
            self.save_hyperparameters()
            self.rnn = nn.RNN(num_inputs, num_hiddens)

        def forward(self, inputs, H=None):
            return self.rnn(inputs, H)

```

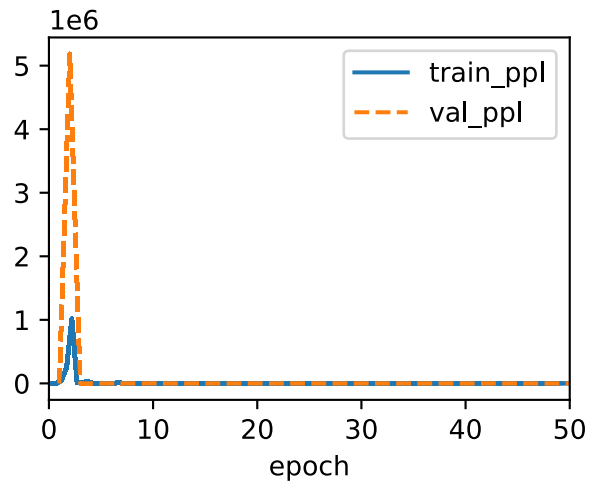
The comparison:

```

[34]: # RNN
rnn = RNN(num_inputs=len(data.vocab), num_hiddens=64)
model = d2l.RNNLM(rnn, vocab_size=len(data.vocab), lr=4)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())

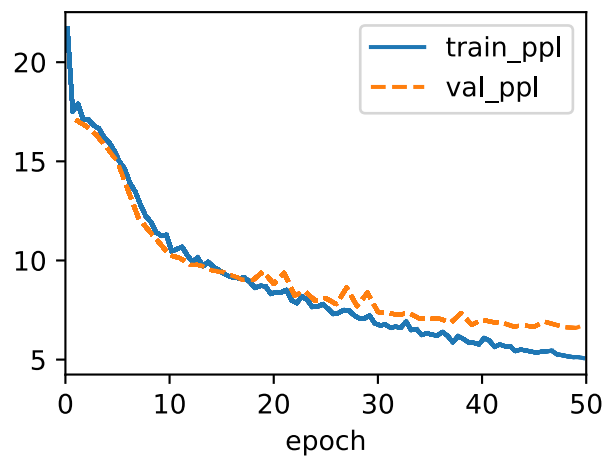
```

[34]: 'it hasoanoanoanoanoanoa'



```
[35]: # LSTM
lstm = LSTM(num_inputs=len(data.vocab), num_hidden=64)
model = d2l.RNNLM(lstm, vocab_size=len(data.vocab), lr=4)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

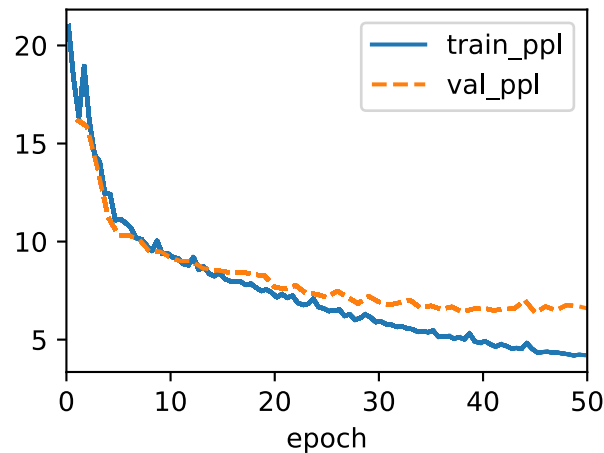
[35]: 'it has i the time travelle'



```
[36]: # GRU
gru = GRU(num_inputs=len(data.vocab), num_hidden=64)
model = d2l.RNNLM(gru, vocab_size=len(data.vocab), lr=4)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
```

```
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

[36]: 'it has the the the the the'



1.2 Problem 2

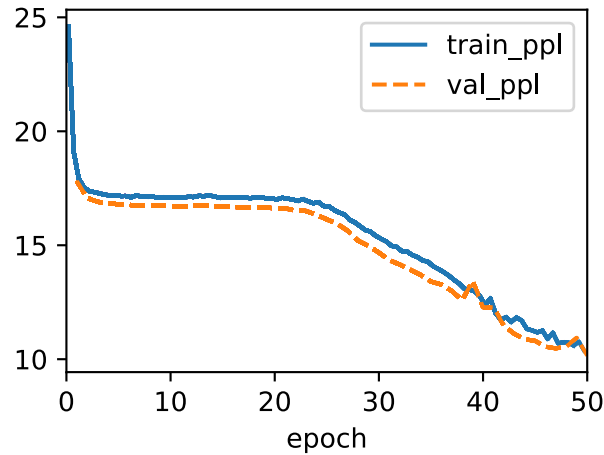
This homework focuses on the Deep RNN problem we did in the lectures.

1. Build the model by replacing the GRU with an LSTM and compare the training and validation loss, and the output sequence (try a few examples) against GRU.

```
[38]: class LSTM(d2l.RNN):
    def __init__(self, num_inputs, num_hiddens, num_layers, dropout=0):
        d2l.Module.__init__(self)
        self.save_hyperparameters()
        self.rnn = nn.LSTM(num_inputs, num_hiddens, num_layers,
                           dropout=dropout)
```

```
[39]: lstm = LSTM(num_inputs=len(data.vocab), num_hiddens=32, num_layers=2)
model = d2l.RNNLM(lstm, vocab_size=len(data.vocab), lr=2)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

[39]: 'it has the the the the the'

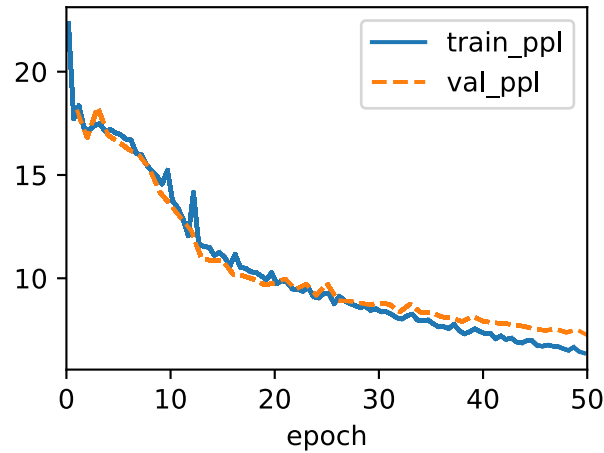


2. Compare runtime for training and inference, computational and mode size complexities, and the output strings for `nn.LSTM` and `rnn.GRU` implementations with each other.

```
[40]: class GRU(d2l.RNN):
    def __init__(self, num_inputs, num_hiddens, num_layers, dropout=0):
        d2l.Module.__init__(self)
        self.save_hyperparameters()
        self.rnn = nn.GRU(num_inputs, num_hiddens, num_layers,
                           dropout=dropout)
```

```
[41]: gru = GRU(num_inputs=len(data.vocab), num_hiddens=32, num_layers=2)
model = d2l.RNNLM(gru, vocab_size=len(data.vocab), lr=2)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

```
[41]: 'it has the the time t'
```

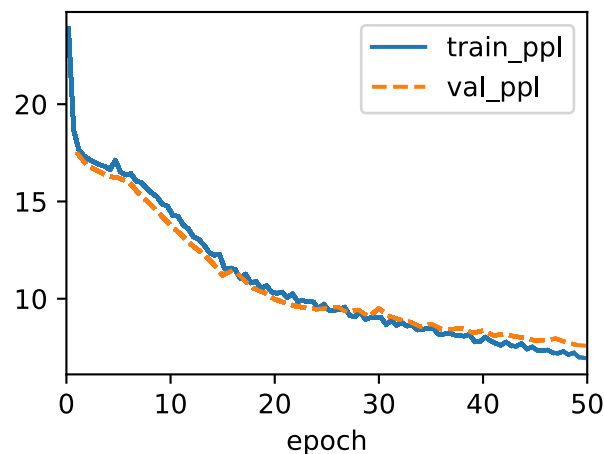


3. Adjust the hyperparameters (fully connected network, number of hidden layers, and the number of hidden states) and compare your results (training and validation loss, computation complexity, model size, training and inference time, and the output sequence). Analyze their influence on accuracy, running time, and computational perplexity.

Adjustment of LSTM Adjustment 1:

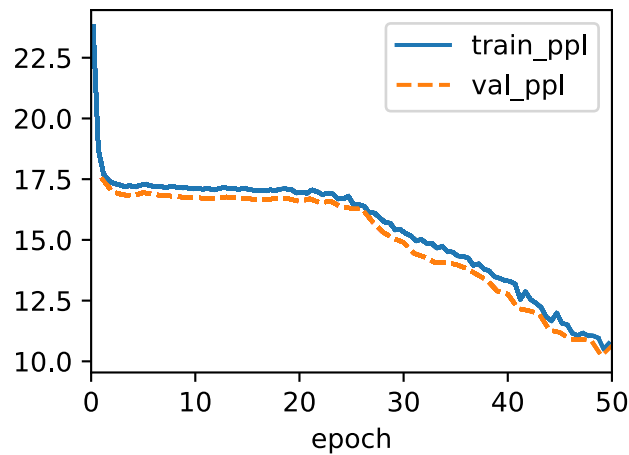
```
[42]: lstm = LSTM(num_inputs=len(data.vocab), num_hiddens=64, num_layers=1)
model = d2l.RNNLM(lstm, vocab_size=len(data.vocab), lr=2)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

[42]: 'it has is aller the the th'



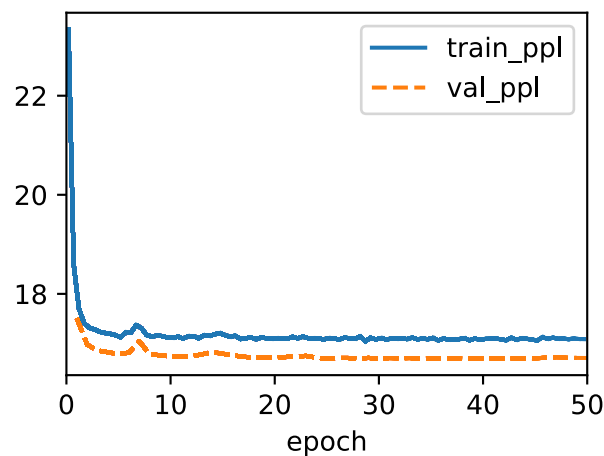
```
[43]: lstm = LSTM(num_inputs=len(data.vocab), num_hiddens=64, num_layers=2)
model = d2l.RNNLM(lstm, vocab_size=len(data.vocab), lr=2)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

[43]: 'it has the the the the the'



```
[44]: lstm = LSTM(num_inputs=len(data.vocab), num_hiddens=64, num_layers=4)
model = d2l.RNNLM(lstm, vocab_size=len(data.vocab), lr=2)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

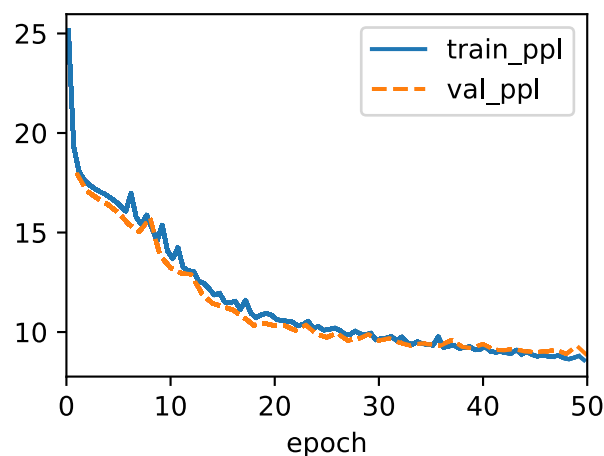
[44]: 'it has'



Adjustment 2:

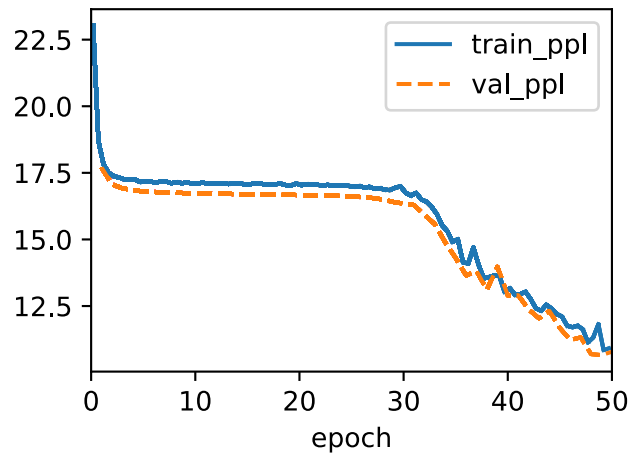
```
[45]: lstm = LSTM(num_inputs=len(data.vocab), num_hidden=12, num_layers=1)
model = d2l.RNNLM(lstm, vocab_size=len(data.vocab), lr=2)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

[45]: 'it has the the the the the'



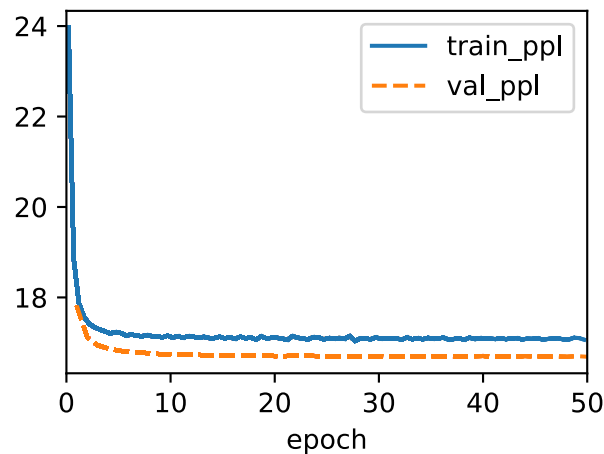
```
[46]: lstm = LSTM(num_inputs=len(data.vocab), num_hidden=12, num_layers=2)
model = d2l.RNNLM(lstm, vocab_size=len(data.vocab), lr=2)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

[46]: 'it has the the the the the'



```
[47]: lstm = LSTM(num_inputs=len(data.vocab), num_hiddens=12, num_layers=4)
model = d2l.RNNLM(lstm, vocab_size=len(data.vocab), lr=2)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

[47]: 'it has'

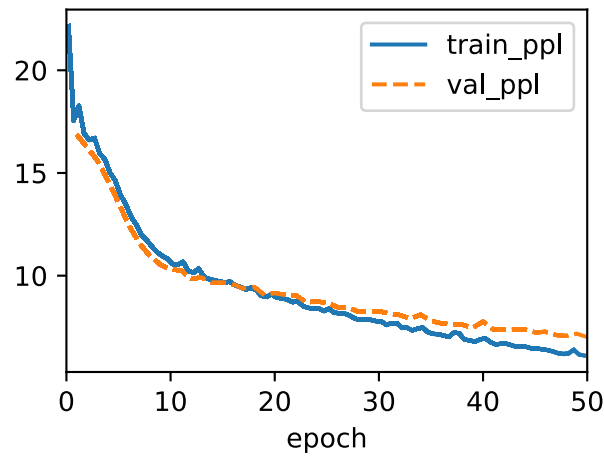


Adjustment of GRU: Adjustment 1:

```
[48]: gru = GRU(num_inputs=len(data.vocab), num_hiddens=64, num_layers=1)
model = d2l.RNNLM(gru, vocab_size=len(data.vocab), lr=2)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
```

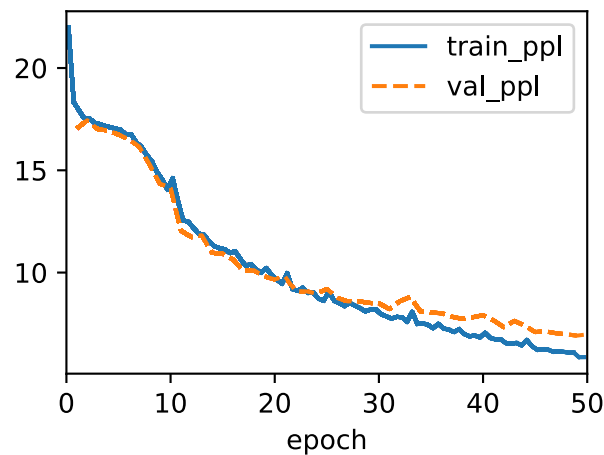
```
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

[48]: 'it has said the the the th'



```
[49]: gru = GRU(num_inputs=len(data.vocab), num_hiddens=64, num_layers=2)
model = d2l.RNNLM(gru, vocab_size=len(data.vocab), lr=2)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

[49]: 'it has so man and and and '



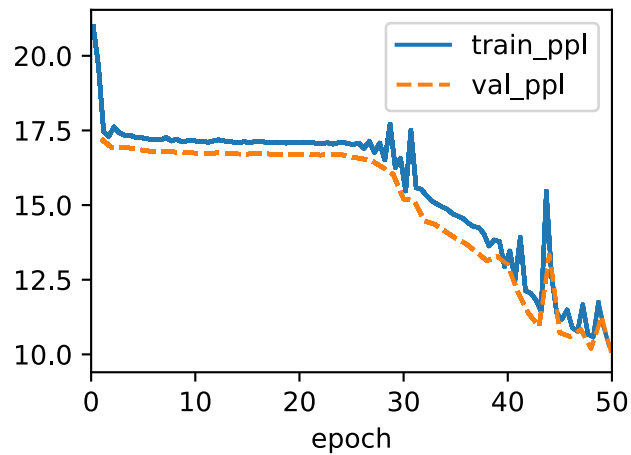
```
[50]: gru = GRU(num_inputs=len(data.vocab), num_hiddens=64, num_layers=4)
model = d2l.RNNLM(gru, vocab_size=len(data.vocab), lr=2)
```

```

trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())

```

[50]: 'it has the the the the the'



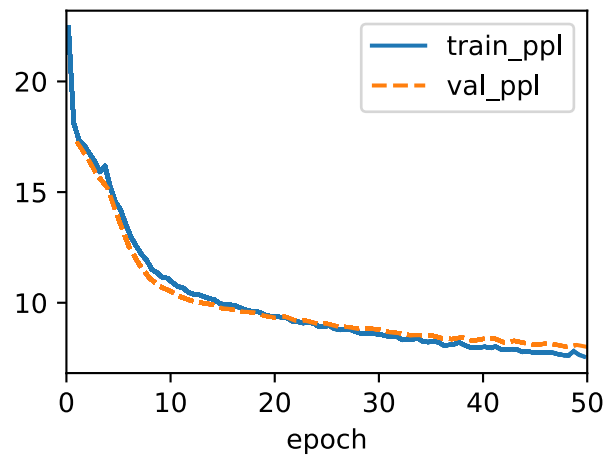
Adjustment 2:

```

[51]: gru = GRU(num_inputs=len(data.vocab), num_hiddens=12, num_layers=1)
model = d2l.RNNLM(gru, vocab_size=len(data.vocab), lr=2)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())

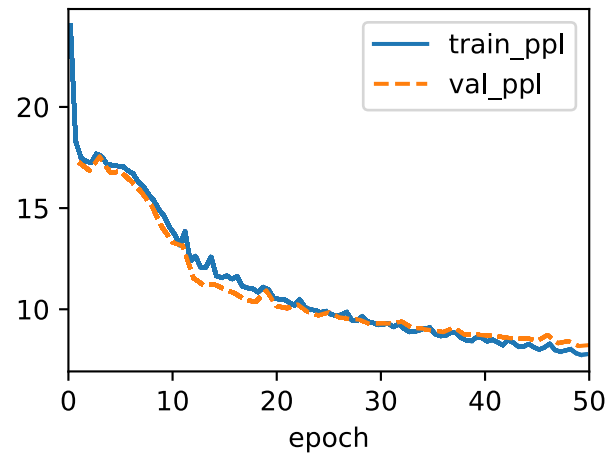
```

[51]: 'it has the the the the the'



```
[52]: gru = GRU(num_inputs=len(data.vocab), num_hidden=12, num_layers=2)
model = d2l.RNNLM(gru, vocab_size=len(data.vocab), lr=2)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

[52]: 'it has the the the the the'



```
[53]: gru = GRU(num_inputs=len(data.vocab), num_hidden=12, num_layers=4)
model = d2l.RNNLM(gru, vocab_size=len(data.vocab), lr=2)
trainer = d2l.Trainer(max_epochs=50, gradient_clip_val=1, num_gpus=1)
trainer.fit(model, data)
model.predict("it has", 20, data.vocab, d2l.try_gpu())
```

[53]: 'it has a a a e a e'

